Massachusetts Institute of Technology
Artificial Intelligence Laboratory

Working Paper 128 May 1976

On the Representation and Use of Semantic Categories: A Survey and Prospectus

Bruce R. Schatz

Abstract:

This paper is intended as a brief introduction to several issues concerning semantic categories. These are the everyday, factual groupings of world knowledge according to some similarity in characteristics. Some psychological data concerning the structure, formation, and use of categories is surveyed. Then several psychological models (set-theoretic and network) are considered. Various artificial intelligence representations (concerning the symbol mapping and recognition problems) dealing with similar issues are also reviewed. It is argued that these data and representations approach semantic categories at too abstract a level and a set of guidelines which may be helpful in constructing a microworld are given.

This report describes research conducted at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the Laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract number N00014-75-C-0643.

"The reason why cognitive processes have for so long been treated as a system of diffuse associative reproductions is that the essential difference has not been appreciated between an associative system and the specific reaction systems which characterize biological processes." - O. Selz (1940)

Introduction

The information processing problem under discussion here is that of semantic categories. These are a crucial piece of semantic memory - the set of facts and organized knowledge needed for cognition and language (as opposed to the experiential knowledge stored in episodic memory - see Tulving(1972)). These categories consist of groupings of simple, factual world knowledge according to some similarity in characteristics. For example, a robin is a member of the category bird. Only generalized members of categories are considered so that one could refer to a robin being a bird but not to Fred being a robin. These categories have the property that they depend crucially on a person's knowledge base so that the actual groupings are individual and domain dependent. Knowledge of sequences such as typical "scripts" (Schank(1975)) are not included. "Category" in this paper will refer to semantic categories.

There are various psychological results dealing with the formation, usage, and types of categories. There are also various mechanisms for representing and using hierarchies (the symbol-mapping and recognition problems) which, although proposed by researchers in artificial intelligence primarily to facilitate problem solving, seem to deal with similiar questions. Most of these approaches deal with categories at a fairly abstract level which appears to be too general to capture some of the most interesting aspects of the problem. Different types of categories seem to have differing structures and usage, distinctions not captured very precisely in these general descriptions which try for universality.

This paper will summarize some relevant psychological data, psychological models, and artificial intelligence representations relating to semantic categories. With the aim of providing a computational model, a sample set of constraints on special cases to consider will be given. It is hoped that this will be useful in choosing a microworld in which to investigate semantic categories (both in laying out what needs to be explained and in describing what criteria a solution might meet). An attempt has been made to present an accurate account of the various ideas considered and to leave applications and interpretations to the reader.

The representation for semantic categories must be complex enough to contain the relevant hierarchies (e.g. a robin is a bird is an animal is a living-thing is a physical object) yet be compact enough to be able to find what is needed or to jump into the appropriate context.

The use breaks down into two major subproblems - property inheritance and recognition. Property inheritance (induction) means that the properties of a category member default to those of the category (although there can be specific exceptions). For example, birds breathe and robins are birds so robins breathe. The problem here is the tradeoff between automatically storing all possible inheritances and deducing them at each stage when necessary.

Recognition entails matching a given description to the appropriate category member. There are thus two stages: finding a set of properties from the description (eg analyzing a visual image), and matching a member to those properties (via intersection or pattern matching). Semantic categories deal primarily with the latter problem. The difficulty there is restricting the matching sufficiently to prevent endless searching and multiple matches.

Thus we are looking for a model and representation which
(1) effectively and efficiently performs the above functions by using context and other information to cut down on deduction time and database search.
(2) accounts for the solid psychological data (particularly with regard to known constraints on structure and formation).
(3) works well in most cases and fails gracefully in the others.
(4) accounts for incomplete knowledge and exceptions (inconsistent data).

The categories considered here will be natural (eg not artificial nonsense words), primarily of concrete objects, and primarily of a semantic or cognitive, not perceptual or physiological nature (eg animal or furniture not color or phoneme).

Psychological data on semantic categories

There are several issues to consider including formation, types of categories, hierarchical structure, and variations in property inheritance. Several of the most useful limiting results (those which place constraints on what a solution must look like) are discussed here. These results do not hold for all semantic categories, but only for those subclasses considered (predominantly categories of concrete objects). Note the general nature of objects considered, eg one might discuss robins but never a particular robin.

(a) *Categories are formed by family resemblances (not by strict defining criteria). Prototypicality of a member is proportional to its degree of family resemblance.*

When attempting to decide whether X is a member of category Y, there are several methods that spring to mind. The simplest and one assumed by many people (Winston(1975), for example)) is a all or none set of criteria. For instance, if X has these 10 properties it is in Y, otherwise it is not. This is unsatisfactory for semantic categories, as a bird can fly but a penguin is a bird and cannot, and a cow with 3 legs is still a cow. An alternate approach was proposed by Wittgenstein(1953) - that of a family resemblance. Here the various members have some properties in common with each other but few or no properties are common to all. Rosch and Mervis(1975) ran a series of experiments that indicated that this type of formation was present in categories of concrete objects and that the degree of family resemblance was proportional to the prototypicality of the member. In addition, membership in another category was inversely proportional to prototypicality. So for example, a car is a prototypical vehicle and thus its properties are fairly common among vehicles while a pushcart is not a very typical vehicle and in fact might even be furniture.

Rosch and Mervis took six common concrete categories (furniture, vehicle, fruit, vegetable, weapon, clothing) and listed 20 members of each. They showed one member of each to 400 students and gave them a minute and a half to list as many attributes as possible. Then the attributes were counted and for each category an attribute was given a score of 1-20 based on how often it had occured. For two of the categories no attribute always occurred (score of 20) and for the other four there was only one such (and for fruit and vegetable this was "you eat it" which is common to other categories). The weightings for the attributes were fairly evenly distributed between 5 and 15 with somewhat more less than 5 and somewhat fewer greater than 15. So a family resemblance, not a criterial check, seemed to be used.

Prototypicality ratings were obtained for the members in each category by asking that a member be rated on a scale of 1-7 as to how well the item fit the subject's idea of the meaning of the category name. These prototypicality ratings were correlated with the degree of family resemblance (computed by summing the weights of all attributes of a member of a category). All of the correlations were high (> .85).

There was also a check for category dominance vs. prototypicality. The former was computed by asking subjects to list 3 categories to which a given item belonged and measuring (freq of the most common category - freq of 2nd most common) + (freq of most common - freq of 3rd most common). Category dominance was strongly correlated with prototypicality. For example, with car, the prototypical furniture, furniture would always be listed first while with telephone, a not very typical furniture, the placing of furniture would vary.

Similar results held for basic categories (see below) and artificial categories (nonsense words). Note that for basic categories, actual contrasting categories (from two different superordinate categories (eg car,chair)) could be used to show that the more prototypical a category member is, the fewer attributes it shares with members of contrasting categories.

N.B. Note that defining criteria can hold for other types of categories, body parts for example.

It might be thought that frequency alone determined the prototype (ie the most common member would be the prototype). However, Rosch, Simpson, and Miller (in press) have shown that for several types of categories, structure seems more important. In one test, subjects were shown dot patterns (Posner dots) generated by a random walk from a set pattern. The subjects learned four such patterns (which did not include the generator) then were given groups of these patterns (including the generator) and asked to rate them according to typicality. The generator was judged to be the prototype despite the fact that its frequency was the least (ie 0).

(b) *There are at least three distinct levels of (concrete) categories: superordinate, basic, subordinate. Basic categories are the level at which the most information is gained about a member.*

It is an empirically determined fact that there are several types of categories depending on generality and how many attributes the members have in common (Rosch and Mervis(1975), Rosch(in press)). (Attributes here will refer to any commonly agreed-upon property of an object.)

Superordinate categories are the abstract ones - different members of this type have overlapping attributes with few attributes common to all members. There are a (small) finite number of members.

Basic categories are the level at which the cue validity is highest (the conditional probability that cue(property) X predicts category Y). Any higher level has fewer attributes in common among members while any lower level has more attributes in common with contrasting categories. There are an infinite number of members.

Subordinate categories contain little more general information than their basic categories but do contain some additional specific information. There are an infinite number of members.

An example of a hierarchy (superordinate, basic, subordinate) of categories would be vehicle, car, sports car. There are a limited number of kinds of vehicles and these do not share many properties. On the contrary, there are many types of cars and these share a number of common properties. Further elaboration, eg sports car, does not give much additional information for usual purposes. Another example is furniture, chair, kitchen chair.

Basic categories are of special interest as they are the fundamental level at which abstractions are made upon the world. Rosch, et.al. (in press) ran a series of experiments to attempt to discover whether there was a most inclusive level above which further classifications yielded much less information and below which further classifications yielded little additional information. A typical experiment went as follows. Nine common (several members with high word frequency) superordinate categories of concrete objects were selected after eliminating possibilities whose members always bore a parts/whole relationship (eg body) or who cut across many taxonomic boundaries (eg food). Then typical basic (ones with unambigious superordinates and several subordinates) and subordinate members were generated, and subjects were given a minute and a half to list all attributes of various presented objects. Sample category hierarchies were:

taxonomy	superordinate	basic	subordinates
non-biological	furniture	table	kitchen table, dining room table
		lamp	floor lamp, desk lamp
		chair	kitchen chair, living room chair
biological	bird	cardinal	Easter cardinal, grey-tailed cardinal
		eagle	bald eagle, golden eagle
		sparrow	song sparrow, field sparrow

biological superordinates: tree,fish,bird
non-biological: musical instrument,tool,clothing,furniture,vehicle,fruit

The results of this and of a subsequent test with subjects judging the correctness of the responses were quite clear, namely for non-biological taxonomies, the members of superordinate categories had few attributes in common while members of basic categories had significantly more, and members of subordinate categories had approximately the same number as the basic ones. For example, the number of common attributes for the judged attributes of furniture was: superordinate 0 basic 7.0 subordinate 7.8. For the biological taxonomies, however, the level of highest cue validity seemed to be that of the superordinate category (eg bird or fish). This can perhaps be explained by the region specificity of biological categories. Berlin (1972) has considerable evidence that the natural grouping (at least for plants) is at the genus level. However, the many species of two genuses (eg maples and oaks) can overlap considerably in their properties and the prototype depends largely on the region considered. So the basic cut for biological categories is somewhat higher (at the tree, fish, or bird level). In addition, anomolous (against the prototype) situations can be dealt with by splitting the categories (eg considering birds as flying birds or non-flying birds at this higher basic level to handle penguins and ostriches.) Another possible explanation is that city dwellers who are not biologists may not make very finely detailed categorizations of biological categories and thus miss the lower level structure.

This leads to the problem that the degree of expertise was not selected for in the experiments of Rosch, et.al. Since a botanist would have a much more detailed biological taxonomy than the average person, his basic categories would be much lower. In addition, expertise depends on the particular individual, subject, and context. So although basic categories probably do universally exist, their contents are quite variable from person to person.

Similar experiments to the above on names were run on motor movements and shapes (pictures and line drawings) with similar results.

Another set of experiments tested whether basic objects were special in perception and development. Pictures of the categories used above were shown following speaking of a category name (priming), and subjects were asked to (a) select which of two pictures was in the spoken category, or (b) decide whether two pictures were physically identical (after the priming). In both cases, basic category names produced the fastest reaction times (better than superordinate, same as subordinate). When doing object recognition, (ie is this picture in the cued category?), basic category cues produced faster reactions than either superordinate or subordinate (for both true and false examples). Superordinate cues were faster than subordinate on true cues while they were the same on false ones. Some developmental experiments seemed to imply that basic categories were the first sorted and named by children. Finally, free-associating the names of various pictures seemed to produce predominantly the basic level names.

(Collins and Quillian(1972) suggest that there is also a difference in method of learning members for different levels of categories, superordinate learning (eg bird) being done by generalization, but basic level learning (eg canary) being done by discrimination (from other basic members).)

Thus there seem to be several levels of categories of which the basic level (if somewhat context-dependent) yields the most information about the object.

(c) *There is a hierarchical intercategory structure which seems to be organized top-down.*

Warrington(1975) ran tests on three patients with visual object agnosia (inability to recognize common objects but fairly unimpaired cognitive and sensory functions). The deficits appeared strongly to be in semantic memory as the patients' perceptual functioning (both acoustic and visual) was normal. Their knowledge of superordinate categories (tested by asking for meanings of words) was significantly less impaired than their knowledge of subordinate categories. (Here subordinate has a somewhat different meaning than with Rosch. For example, superordinate - animal,plant,inanimate object ; subordinate - color,number of legs). In tests where they were asked to recognize pictures of objects, they could recognize that something was a flower, but not which one. For example, shown a daffodil, they could say it was a flower but not that it was a daffodil. They could differentiate between fruit and vegetables, but had difficulty identifying particular ones.

Occasionally, they would confuse the member with a similar one from the same category (eg donkey-horse, dog-cat).

A visual recognition test (the patients were shown pictures of animals and queried about various properties of these) indicated that the more specific a property, the worse the performance (properties in order of best to worst remembrance: animal, bird, size, English, name). A possible explanation is that semantic memory is hierarchically structured and that the patients were missing the bottom nodes. To test this conjecture further, an auditory recognition task was run where the names of the animals above were spoken and the properties inquired about. Analogous results were obtained. Thus semantic category memory seems to be hierarchically structured, and top-down in the sense that subordinate categories are less essential than superordinate. (The agnosic patients can still function effectively if in a somewhat limited range.)

There is a great deal of reaction time data dealing with what the semantic hierarchy looks like. Collins and Quillian (see (1972)) in their seminal work assumed a semantic network where traversing each link took time so that the more links, the more time. So, for example, "an oak is a tree" is faster than "an oak is a plant". Strong cognitive economy (a property is stored at the highest possible level, eg branches is stored at tree, not at oak) was assumed, so that "an oak has acorns" is faster than "an oak has branches". However, later investigators showed hat the times were confounded with other variables. For instance, Conrad(1972) showed that the frequency of the words used essentially determined the results. Rips,Shoben, and Smith(1973) showed that the effect is in fact reversed for some superordinate categories, eg "X is an animal" is faster than "X is a mammal". This may imply that there is no strict hierarchy or that the hierarchical structure dpends on the level of category (eg basic,superordinate).

The data concerning the structure of the hierarchy is somewhat unclear, but a few generalizations can be made (Kintsch(1974)). Semantic distance, however defined, plays a crucial role in reaction times for retrieval and recognition. Positive judgements (X is a Y) are facilitated if X is semantically close to Y while negative judgements (X is not a Y) are inhibited if X is semantically close to Y. Also, the larger the category, the longer the time.

(d) *There seems to be a probabilistic property inheritance based on the generality of the member with the known property.*

Rips(1975) told subjects to consider the case of 6 birds (robin, sparrow, duck, geese, hawk, eagle) on a desert island. Then he told them that bird X (the given - one of the 6) has a strange new disease and asks them to estimate the probability that each of the other birds (the targets) has it. He found the probability to be a direct function of the semantic distance (measured by a multidimensional scaling solution of subject's ratings) between (a) the given and the target and (b) the category and the given. In other words, if X has it then so do very similar ones (eg robin-sparrow, eagle-hawk, duck-geese). In addition, the more typical the bird with the disease (robin > duck > hawk), the more likely that more of the others have it. (eg X=robin implies all birds have it, while X=duck perhaps implies only fowls have it). Similar results were found for animals (horse,deer,lion,dog,mouse,pig).

Carey (personal communication) has obtained similar, even stronger results with children. She asked 4 year olds whether various animals (people, mammal, bird, insect, fish, worm) had property X. The curve of property inheritance fell off regularly from 100% to 40% (in the order indicated) for all "animal" properties. (even for bones which is false for insects and worms). So the children seemed to be strongly but probabilistically inheriting from the prototypical animal (people). This was despite the fact that they denied that people were animals. An analogue to Rips' experiment was performed by telling the children that people have spleens. The inheritance followed the 100-40% curve. When they were told that dogs have spleens, however, there was no inheritance at all. So apparently properties are inherited according to the judgement of how general the known instances are. (This may imply a probabilistic is-a link. Note that the judgement of generality is person and context dependent.)

Psychological models

There are two major "processing" models for semantic memory. These could be conveniently called set-theoretic and network models. The set-theoretic model is a standard pattern match, trying to find the minimum "semantic distance" (suitably defined) between the set of properties of the object and the set of properties of various categories. Most current ones tend to be two stage. (Property inheritance and structural considerations are ignored.) The network model is an associationistic hierarchical semantic net where inheritance is done by following along the links and recognition is done by intersection up to some recognized node.

One of the best set-theoretic models is the feature comparison model of Smith, Shoben, and Rips (1974). This consists of two stages where the first is an overall comparison match of all properties of the object and the proposed category. There is a negative threshold (answer false) and a positive threshold (answer true) for quick decisions. If the match is not decisive, a second comparison is done, this time only on defining features. So "a robin is a bird" would be a fast match, while "a penguin is a bird" would require both stages and be considerably slower (as actually observed - see (c) above). This agrees nicely with the family resemblance (continuum) idea of formation of categories. (feature overlap = semantic distance which is proportional to prototypicality). These are hard to handle in the standard network model (requiring probabilistic or varying strength links or additional peculiar nodes). Reversals of timing (eg robin is faster than chicken when deciding if they are birds but slower when deciding if they are animals) are easily handled (by the number of overlapping features) while they contradict the network model. Unfortunately, nothing seems to be said about doing property inheritance. And the major drawback is how to do the first stage comparison since there are an infinite number of possible features. Rips, Shoben, and Smith(1973) attempt to overcome this problem by comparing only on "functional" features but deciding what these are appears extremely difficult. In addition, defining sets of characteristics do not seem to exist (eg a cow with 3 legs is still a cow) nor are people always aware of them (what are those of a sponge in answering "is a sponge an animal?"). People tend to be uncertain even of ordinary characteristics (does a sponge breathe?).

(Kintsch(1974) has a similar two stage model where the first stage compares against known relations and the second against inferred ones.)

The network models (see Collins and Loftus (1975)) consist of a semantic network with varying types of links and nodes and processing proceeding by searching through this tree leaving an activation trace which makes the links stronger (thus accomplishing contexting). (Note that set-theoretic models, which do not specify a memory structure, could be implemented in semantic networks. However, the natural processing strategies for the two types of models are quite distinct as are the methods for predicting reaction times.) Collins and Loftus claim different strengths of links (criteralities) so that family resemblances can be simulated. They assume weak cognitive economy, ie if a property of a node is learned which is also a property of an inferior node, the inferior node's property is not erased. However, inferior node properties can still be deduced from their superiors. Thus this model can be patched to account for all of the above data. However, what one is left with appears to be too general and unwieldy to be of interest. There are no special constraints or interesting predictions.

Computational representations

Researchers in artificial intelligence(AI) have proposed a number of representations and processing strategies for dealing with problems similar to those of semantic categories. (Predominantly at a very abstract intuitive level.) Most tend to fall under the rubric of "symbol mapping" (after Fahlman(1975)). A common notation in these Planner-like databases (cf Hewitt(1971)) is the ordered triple of (relation object property) such as (color Clyde gray). This representation appears sufficient for much of semantic category information.

An essential difference between these AI representations and semantic categories is that the latter deal with category members in the abstract while the former are primarily interested in particular instantiations. For example, with semantic categories, one might discover (isa elephant mammal) and want to know what information this gives about elephant. For symbol mapping, however, one is interested in a particular elephant, say Clyde, and wants to know what information (isa Clyde elephant) yields about Clyde. So in some sense, the AI approach is that the database (world knowledge) is fixed and one wants to apply it to actual occurences, while the semantic category approach consists of using (and possibly changing) the database at an abstract level. Various AI methods for utilizing categories will now be described. It should be remembered that the problems being addressed are somewhat different from those of semantic categories, but hopefully the solutions will provide hints for the semantic solutions.

The use of categories, as noted in the introduction, breaks into two pieces (cf Marr and Nishihara(1975)).

(1) *Property inheritance.* Given a member of a category, it inherits its category's properties. For example, (color elephant gray) $\wedge$ (isa Clyde elephant) => (color Clyde gray). Thus properties are deduced from superordinate categories.

(2) *Recognition.* Given an object and/or its properties, find its category. Thus examination (eg visual,acoustic) of the object is done to determine its properties and the properties are intersected to find the category for this member.
There are several types.

- (a) General. Looking for a member. eg gray $\wedge$ trunk $\wedge$ eats-peanuts => elephant.
- (b) Specific. Looking for an instantiation of a member.
  - (b1) Current mention (reference window). eg elephant $\wedge$ male $\wedge$ have just seen Clyde and Martha the elephants => Clyde.
  - (b2) Memory reference (recall). elephant $\wedge$ male => Clyde (whom we saw last week).

Type 2(b) is ruled out by the restriction in this paper to semantic (as opposed to episodic) memory (in the sense that only timeless, "universal" facts are considered here) and by the restriction of semantic categories to only abstract levels of detail. Statements like elephant $\wedge$ flies => Dumbo could still be made but won't be here. An important point to remember when dealing with predominantly isa hierarchies, as those here, is that they tend to be bushy (an elephant has many properties) but not deep (only a few steps, eg an elephant isa herbivore isa mammal isa animal isa living-thing isa physical-object).

Most of the AI effort thus far has been put into property inheritance. That will be primarily discussed with side mentions concerning recognition where relevant. There are two obvious approaches to property inheritance suggested in a Planner-like database; antecedent and consequent reasoning.

Antecedent reasoning is forward chaining. It says if A is asserted, then assert B. So essentially all possible entries are precomputed. For example, when (body-covering mammal hair) becomes known, all mammals will immediately inherit this property (ie (isa X mammal) => (body-covering X hair)). This leads to a tremendous number of relations. No deduction is needed upon retrieval but storage space becomes crucial (This corresponds to no cognitive economy.).

Consequent reasoning is backwards chaining. It says that to prove B, try to prove A. So entries are computed only when needed. For example, to determine (body-covering X ?), all body-coverings relations are examined and their objects checked to see if they are an X (ie (body-covering Y Z) $\wedge$ (isa X Y) => (body-covering X Z)). This leads to a tremendous number of possible deductions. (There are many kinds of body-coverings and the search would be even longer for such relations as color. Note that an object search could also be done - trying all things that are an X and seeing if they have a body-covering property (as opposed to the relation search already described).) Minimal storage is needed but deduction time becomes crucial (This corresponds to strong cognitive economy.).

The antecedent size blowup and the consequent deduction blowup are not hindrances when dealing with a small amount of total world knowledge. This is the case in some domains (see particularly EL (Sussman and Stallman(1975)) - an electronic circuit analyzer which does model-driven antecedent deduction and MYCIN (Davis,Buchanan, and Shortliffe(1975)) - a medical consultation program for bacterial infection which does data-driven consequent reasoning. Both use rule-based mechanisms which are effective because of the limited knowledge necessary.) This is emphatically not the case with the essentially arbitrary semantic categories and thus more sophisticated methods are necessary.

What is needed is to restrict the deduction search tree and the storage space required or both. There are several current methods, some fairly complete (contexting, relation typing, parallelism) and some heuristic. The complete methods, which consider all inputs equally, have the problem that the number of input properties must somehow be kept to a small finite number. The heuristic methods, which consider the inputs selectively (gross to finer characteristics for the example here (Marr and Nishihara)), do not have this problem. But they must then provide a way to make up for their lack of complete information.

McDermott (1975b, 1975a) does a kind of antecedent reasoning in which deductions are postponed on a relation until a request for the object is made (placing it in the current context). Each category (eg elephant) has a packet associated with it. This is a grouping of its properties with the object left unspecified to be filled in when the packet is activated and the object instantiated. For example, (isa elephant ?X) => (color ?X gray) $\wedge$ (isa ?X mammal) $\wedge$ (nose ?X trunk) ... When some elephant (say Clyde) is mentioned, the packet is activated (comes into the current context a la Conniver (McDermott and Sussman(1974))). Property inheritance is now easy if there is only a small number of packets available at any one time. The problem with the contexting approach is that one must insure that the number of activated packets is indeed small. McDermott claims this to be true in his electronic design research, but the assumption seems somewhat unreasonable for real-world semantic categories.

Moore (1975a,1975b) uses the usual Planner pattern matching but adds type checking to cut down on search. Relations are grouped under the category they belong to, eg (color Clyde/elephant gray) or (color ?X/elephant gray), and for each instance a list of its supercategories is made, eg Clyde isa (elephant, herbivore, mammal, animal, living-thing, physical-object). To do property inheritance, one goes through the supercategory list, checking for the property in the hash table (bucket) of each. For example, to deduce (body-covering Clyde ?), one looks through the buckets of elephant,herbivore,mammal, and so on until a body-covering property is found. This method assumes only one type of subsetting relation (namely isa) and will not allow tangled hierarchies (eg (isa Clyde male) and (isa Clyde herbivore) but neither male nor herbivore is a subclass of the other).

Fahlman (1975a,1975b) has proposed using special parallel hardware to do normal consequent reasoning. For example, to deduce (color Clyde ?), mark all categories that contain Clyde and travel down the isa link, checking sideways for a color property. Since this is done in parallel, it makes no difference how many properties or members there are at a level of the hierarchy. In particular, tangled hierarchies are handled as easily as untangled ones as all properties at level n above Clyde are checked simultaneously regardless of how many members they belong to. Other relations can be handled because of this by turning them into isa's, eg (occupation Clyde taxi-driver) into (isa Clyde taxi-driver). This does have the unfortunate side-effect of making isa into merely punctuation (as McDermott(1975b) points out). Recognition (of the type assuming the properties are given) can be easily handled by marking all nodes containing any of the given properties and finding the nearest intersection. The problem here is the vast overkill for doing property inheritance (parallelism replaces knowledge of context) and the reliance on intersection as the way to do recognition (many recognition problems, eg vision, entail intersection only as a small subpiece).

Marr and Nishihara (1975,Marr(1975)) use a heuristic partial solution in their work on visual recognition (functional semantics are not considered here). The recognition is of the property-determining sort in that a model of the properties of an object is built. The objective is to compute a 3-dimensional model from a 2-dimensional projection without actually caring what the object is. The database consists of groupings of properties called templates which are similar to unactivated McDermott packets except that all of the property values are indirect references (eg template = elephant, body-covering = $elephant-body-covering where $ indicates one must look up the value (which can change from time to time).) Only particular instances of templates can be considered so that a template must be instantiated before use. Not all properties are stored in the templates (the system is incomplete), but there exists a global index to determine where to find related relevant information.

For example, assume (isa Clyde elephant) is given and (hair Clyde ?) is asked. (isa Clyde elephant) causes an instantiation of the elephant template, $Clyde, to appear. Initially $Clyde contains no information except class elephant (which means the properties default to those in the elephant template). $Clyde is unsuccessfully searched (via elephant) for a hair property. So the global index is checked for hair packets (property templates relating to hair). $body-covering is found. ($ indicates it must be evaluated via a template to get the current value.) The $Clyde instance is unsuccessfully searched for body-covering. (via the elephant template). Now the elephant template has the property class mammal so a mammal template is instantiated for Clyde. This is unsuccessfully searched for a hair property but the property body-covering here is equal to $mammal-body-covering. So the template mammal-body-covering is checked for the property hair and it returns (hair yes). Thus a new property is created for the elephant Clyde instantiation, namely (body-covering $100) where $100 begins as a pointer to mammal-body-covering. And the question is answered yes.

As noted above, this method has no global search solution to property inheritance but instead uses a global index to find out where to get the necessary information. All references are indirect so exceptions are allowed (Clyde the elephant has a different instantiation from Martha the elephant and any of their properties could differ although the default references are the same). Contexting is built-in by allowing the reference values to change ($mammal-body-covering could be changed to evaluate to skin without affecting the template(s)). The additional indexing necessary for intersection recognition is available by using the global index and by indexing the templates that specify 3-dimensional shapes. This appears quite workable for recognizing objects given at least their coarse descriptions (in reasonable views) although it remains to be seen how many shape templates are necessary to be effective for arbitrary objects. In any case, it is not certain (although it seems reasonable) that extending this scheme to the functional properties needed for semantic categories would be successful.

A few other mechanisms, essentially unrelated to the symbol mapping controversy, should be mentioned here.

Greenblatt(1976) proposes a hierarchical top-down tree (inspired by Warrington - see (c) under psychological data). Each node has a thread to the top of its tree so that the exploration of properties goes from general to specific. For example, in the semantic tree living-thing -> animal -> bird -> duck -> mallard, mallard would have a thread (pointer) to living-thing. To do recognition from a description when a mallard is suspected, the properties of living-thing would be examined then the properties of animal and so on. If the object turned out to be some other duck, the properties examined so far would still be relevant. This strict top-down approach (interpretation of Warrington's data also) appears quite strongly in Marr and Nishihara's method as well.

A linear threshold algorithm on the properties of a member is a possibility. For example, for birds, fly gets 10, beak gets 10, red gets 2, chirps gets 3, and so on. The numbers are summed for the properties observed and the object is recognized as a bird if above a threshold. The problem is where the numbers come from and how to limit the properties to a small finite number. This approach worked well with mildly arbitrary numbers in MYCIN mentioned earlier, but bacterial diseases are a much more restrictive world than real-world semantic categories.

A more promising variant on this is a frame system approach using varying slots (Minsky (1975)). For example, Pauker et al. (1975) in a medical diagnosis program (taking the present illness - asking questions of a patient and attempting to determine his illness) used a simple version which had a set of properties for each disease considered categorized by such quantities as must-not-have, is-sufficient, major-scoring(with numbers), minor-scoring(with numbers). This jumps through various widely varying premises to get a handle on possible illnesses. A category example would be:
name: bird
must-not-have: live-born-young, gills
is-sufficient(usually): has-wings, can-fly, has-a-beak, has-feathers
major-scoring: biped .7, three-toed-feet .6, migrates-south-for-winter .5
minor-scoring: chirps .4, warm-blooded .3, eats worms .2

The problem is that again negative conditions are hard to find and any positive condition could be left out.

The Category Microworld

The recurring problem in considering processing semantic categories is how to determine a reasonable ("defining") set of properties for each member and yet avoid being stuck by exceptions. Physicians avoid these difficulties by defining diseases in an approximately all or none fashion (based on physiological or anatomical conditions - note that medical diagnosis usually consists of attempting to guess the pathology from outward symptoms or from safe, easily run clinical tests.)

Biological taxonomists behave similarly. Taxonomic boundaries are drawn by such characteristics as cell structure, biochemistry, reproduction,distribution (ecology), phylogeny(evolution), movement, and morphological diversity. Objects differing in these characteristics are placed in different slots in the hierarchy. Semantic categories represent everyday taxonomic distinctions which, unfortunately for the prospective theorist, are informal and somewhat inconsistent. Perhaps the best (and possibly the only) way to investigate these is to consider their behavior in some simple subsystem or microworld.

One possible contender is to imitate some small subsection of ("essentially solved") biological taxonomy on a psychological (everyday, folk usage) level. Much is known about the members and their properties (both macro and microscopic) - see for example Scagel,et. al. (1965). Thus writing a program to recognize some semantic subclass of biological knowledge (eg plants) might be fruitful. This has been reasonably done already for biological taxonomies (Morse (1974) has a series of programs which, given a formal description of a wide variety of plants, searches the decision trees to produce the taxonomic classification.) and some quite effective numerical techniques for classifying such taxonomies are also known (Jardine and Sibson(1971) for example). So some techniques could possibly be borrowed.

Choosing a reasonable microworld is a tricky business. Based on observations of above mentioned psychological data and artificial representations, a sample set of rules which may be helpful in constructing one to consider the computational use and representation of semantic categories will now be listed.

(1) The category is already known (no formation or learning).
(2) The current state of knowledge is sufficient. (eg If it is not known that a robin is a thrush, that is all right. Incomplete knowledge always holds.)
(3) The situation must have and use prototypes.
(4) Basic level categories plus some hierarchy should be included.
(5) Either some type of recognition or property inheritance but probably not both should be considered.
(6) Consider semantic not episodic memory.
(7) Limited subsetting relations should be allowed (eg only a true isa).
(8) Pick a particular task (say visual not functional recognition).
(9) Consider general instances not specifics (to avoid the recall problem. cf (6)).

Even with these caveats, this processing problem may still be intractable.

Conclusion

A survey has been given here of various psychological data and computational representations concerning semantic categories at a rather abstract level. The problems with the nebulous nature of these categories seems to indicate that attempts to provide a general solution will probably meet with little success. Different types of categories require different solutions. For example, phonological categories do not seem to have the prototype structure that the concrete categories mentioned here do. (One can not decide which instance of the phoneme ba is more typical and in fact cannot differentiate between two instances of ba even if spoken at different frequencies. So apparently a criterial definition does hold here (Fodor,Bever, and Garrett(1974).) Another example is that there is some evidence that the prototypes of physiological categories (eg red) are cross-culturally universal (Heider (1972)) whereas this is certainly not the case for biological categories (eg birds,trees) whose prototypes vary from region to region and individual to individual. In fact, one can probably make a case that there are different processors for property inheritance and recognition, for retrieval and storage, and for superordinate, basic, and subordinate categories as well as numerous special purpose types.

Explaining some of the general psychological data discussed above would be of interest if done fairly precisely in some restricted context. It is hoped that the descriptions here of some artificial intelligence representations and some possible criteria for deciding on a micro-world will be of help to future explorers of semantic categories.

Acknowledgements

Prof. Susan Carey and Dr. David Marr provided much of the inspiration for the philosophy stated here. Prof. Carey, in addition, provided much help in the author's attempts to understand pieces of cognitive psychology. Drew McDermott performed a similar function for the artificial intelligence portions. Both the content and clarity of this paper were greatly improved by criticism from Tomas Lozano-Perez, David Marr, Drew McDermott, and Jan Galkowski.

Bibliography

*** indicates particularly good introduction to the psychological data
*** indicates the same for the artificial intelligence representations

Berlin, B. Speculations on the growth of ethnobiological nomenclature. *Language in Society* 1:51-86(1972).

Carey, S. personal communication and MIT course 9.65 Cognitive Processes (spring 1976).

Collins, A.M. and Loftus, E.F. A spreading activation theory of semantic processing. *Psychological Review* 82:407-428(1975).

*** Collins, A.M. and Quillian, M.R. How to make a language user. Chapter 8 (pp310-351) in Organization of Memory ed by E. Tulving and W. Donaldson (New York: Academic Press, 1972).

Davis, R., Buchanan, B., and Shortliffe, E. Production rules as a representation for a knowledge-based consultation program. Stanford Artificial Intelligence Lab Memo AIM-266 (Oct 1975).

Fahlman, S.E. Thesis progress report: A system for representating and using real-world knowledge. MIT Artificial Intelligence Lab Memo 331 (May 1975). (a)

Fahlman, S.E. Symbol mapping and frames. *ACM SIGART Newsletter* 7-8 (Aug 1975). (b)

Fodor, J.A., Bever, T.G., and Garrett, M.F. The Psychology of Language. (New York: McGraw-Hill, 1974).

Greenblatt, R. Top-down semantic memory. unpublished paper (April 1976).

Heider, E. R. Universals in color naming and memory. *J. Experimental Psychology* 93:10-20 (1972).

Hewitt, C. Procedural embedding of knowledge in PLANNER. pp167-182 in Proc Second International Joint Conf. on Artificial Intellgence. (London: The British Computer Society, Sept 1971).

Jardine, N. and Sibson, R. Mathematical Taxonomy. (London:Wiley & Sons,1971).

*** Kintsch, W. Memory Search II: The use of knowledge in the verification of statements. Chapter 10 in Representation of Meaning in Memory (pp195-218) (Potomac,Maryland: Lawrence Erlbaum Associates, 1974).

Marr, D. Less about symbol-mapping. *ACM SIGART Newsletter* 3-4 (Dec 1975).

*** Marr, D. and Nishihara, H.K. Spatial disposition of axes in a generalized cylinder representation of objects that do not encompass the viewer. MIT Artificial Intelligence Lab Memo 341 (Dec 1975).

McDermott, D.V. and Sussman, G.J. The CONNIVER Reference Manual. MIT Artificial Intelligence Lab Memo 259a (Jan 1974).

McDermott, D.V. Symbol-mapping:a technical problem in PLANNER-like systems, *ACM SIGART Newsletter* 4-5(April 1975). A packet-based approach to the symbol-mapping problem, *ACM SIGART Newsletter* 6-7(Aug 1975). (a)

*** McDermott, D.V. Very large PLANNER-like data bases. MIT Artificial Intelligence Lab Memo 339 (Sept 1975). (b)

Minsky, M. A framework for representing knowledge. p211-277 in Psychology of Computer Vision ed by P.H. Winston, (New York: McGraw-Hill,1975).

Moore, R.C. A serial scheme for the inheritance of properties. *ACM SIGART Newsletter* 8-9 (Aug 1975). (a)

Moore, R.C. Reasoning from incomplete knowledge in a procedural deduction system. MIT Artificial Intelligence Lab TR-347 (Dec 1975). (b)

Morse, L.J. Computer Programs for Specimen Identification. Biological Series, Vol 5, No 1, (East Lansing: Publications of the Museum, Michigan State University).

Pauker, S.G., Gorry, G.A., Kassirer, J.P., and Schwartz, W.B. Towards the simulation of clinical cognition: Taking a present illness by computer. *American J. of Medicine* (in press - 1975).

Rips, L.J. Inductive judgements about natural categories. *J. Verbal Learning and Verbal Behavior* 14:665-681(1975).

Rips, L.J., Shoben, E.J., and Smith, E.E. Semantic distance and the verification of semantic relations. *J. Verbal Learning and Verbal Behavior* 12: 1-20 (1973).

Rosch, E. Cognitive reference points. *Cognitive Psychology* 7:532-547(1975).

*** Rosch, E. Classifications of real-world objects: origins and representations in cognition. *Bulletin de Psychologie* in press (special issue on semantic memory).

Rosch, E. and Mervis, C.B. Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology* 7:573-605(1975).

Rosch, E., Simpson, C., and Miller, R.S. Structural bases of typicality effects. *J. Experimental Psychology: Human Perception and Performance* (in press).

Rosch, E., et.al. Basic objects in natural categories. *Cognitive Psychology* (in press).

Scagel, R., et.al. An Evolutionary Survey of the Plant Kingdom. (Belmont,Cal.: Wadsworth Publishing Co.,1965).

Schank, R.C. Using knowledge to understand. pp117-121 in Theoretical Issues in Natural Language Processing, ed by R.C. Schank and B.L. Nash-Webber, (MIT, June 1975).

Selz, O. Lecture 8 (1940). Reprinted and translated in Kintsch(1974), p 41.

Smith, E.E., Shoben, E.J., and Rips, L.J. Structure and process in semantic memory: a featural model for semantic decisions. *Psychological Review* 81:214-241(1974).

Sussman, G.J. and Stallman, R.M. Heuristic techniques in computer-aided circuit circuit analysis. *IEEE Transactions on Circuits and Systems* CAS-22: 857-865 (Nov 1975).

Tulving, E. Episodic and semantic memory. Chapter 10 (pp 381-403) in Organization of Memory ed by E. Tulving and W. Donaldson (New York: Academic Press, 1972).

Warrington, E.K. The selective impairment of semantic memory. *Quarterly Journal of Experimental Psychology* 27:635-657 (1975)

Winston, P.H. Learning structural descriptions from examples. pp157-209 in The Psychology of Computer Vision ed by P.H. Winston (New York: McGraw-Hill, 1975).

Wittgenstein, L. Philosophical Investigations. (New York: Macmillan, 1953).